"If you'd wiggled A, then B would've changed"
Causality and counterfactual conditionals

Katrin Schulz[†]

ILLC
University of Amsterdam
k.schulz@uva.nl

Abstract. These are the proofs of the facts of the paper mentioned above. They did not appear in paper itself, because it had to be cut down in length substantially. Because of an tremendous workload the last two years I took me until june 2012 to finish the editing of this part. My apologies!

The proofs themselves are not exciting (we are working with a finite propositional language), but sometimes tiresome in spell out, because they mostly are about checking properties of certain constructions: the immediate consequence operator $\mathcal{T}_\mathcal{D}$ and the basis $B_D(w_0)$ of a possible world w_0.

1. Appendix

DEFINITION 9. *For two sitations s_1, s_2 we define a partial order $\leq$ over situations as follows: $s_1 \leq s_2$ iff for all proposition letters p $s_1(p) \leq s_2(p)$.*

FACT 2. *Let D be a dynamics. The immediate consequence operator $\mathcal{T}_D$ is 'monotone' in the sense that for all situations s it holds $s \leq \mathcal{T}_D(s)$.*

Proof of fact 1: We show that $\forall p \in \mathcal{P} : s(p) = 1 \Rightarrow \mathcal{T}_D(s)(p) = 1$. The proof of the case $\forall p \in \mathcal{P} : s(p) = 0 \Rightarrow \mathcal{T}_D(s)(p) = 0$ works along the same lines.
We assume $s(p) = 1$ for some $p \in \mathcal{P}$.

Case 1 Assume $p \in B$. It follows from definition 3, case (i) that $\mathcal{T}_D(s)(p) = s(p) = 1$.

Case 2 Assume $p \in \mathcal{P} - B$. It follows from definition 3, case (ii.b) that $\mathcal{T}_D(s)(p) = s(p) = 1$.

FACT 3. *Let D be a dynamics and s a situation. There exists a fixed point s^* of the operator $\mathcal{T}_D$ which is the least fixed point with the property $s \leq s^*$. This fixed point is reached after finitely many steps.*

[†] The author is supported by the Netherlands Organization for Scientific Research.

Proof of fact 3: Let $D = \langle B, F \rangle$ be a dynamics and s a situation. We define a sequence of situations $s_0, s_1, s_2, ...$ inductively as follows: $s_0 := s$, $s_{n+1} := \mathcal{T}_D(s_n)$. Because the number of models for a finite propositional vocabulary is finite, we know that there are only finitely many different models in this sequence. Because of fact 2, we also know that for all natural numbers n: $s_n \leq s_{n+1}$. Together, this means that for some n with $\mathcal{T}_D(s_n) = s_n$. Let i be the least natural number with this property. Then s_i is the least fixed point s^* we where looking for. It is reached in i steps.

FACT 4. *(Fact 1 from the main paper)*
For a possible world w_0 and a dynamics D the basis $B_D(w_0)$ exists and is uniquely defined.

Proof of fact 4: Let $D = \langle B, F \rangle$ be a dynamics and w_0 a possible worlds.
Part A We first prove the existence of the basis by construction.

DEFINITION 10. *The contruction of a basis*
For a possible world w_0 and a dynamics D let X be the following set of literals:
(i) If $p \in B$, then if $w_0(p) = 1$ then $p \in X$, otherwise $\neg p \in X$.
(ii) If $p \in \mathcal{P} - B$ with $Z_p = \langle p_1, ..., p_l \rangle$ and $f_p(w_0(p_1), ..., w_0(p_l)) \neq w_0(p)$, then if $w_0(p) = 1$ then $p \in X$, otherwise $\neg p \in X$.
(iii) nothing else is in X.

With fact 3 we know that s^*_X exists. We will show that X is a basis of w_0, i.e. (i) $s^*_X = w_0$ and (ii) X is a minimal set with this property.

Add i). We define recursively for all $p \in \mathcal{P}$ the hight of p, $H(p)$ to be the maximal distance of p from a background variable.
(i) If $p \in B$ then $H(p) = 0$.
(ii) If $p \in \mathcal{P} - B$ with $Z_p = \langle p_1, .., p_l \rangle$, then $H(p) = Max\{H(p_i) \mid 1 \leq i \leq l\}$.
We will show that $\forall n \in \mathbb{N} \forall p \in \mathcal{P}[H(p) = n \rightarrow s^*_X(p) = w_0(p)]$ with induction over H. This works because $\mathcal{P}$ is finite and $\langle \mathcal{P}, R^T_F \rangle$ is a poset. So, every p will have a well-defined hight $n \in \mathbb{N}$.
The basis case. If $H(p) = 0$, then $p \in B$. From the definition of X and the definition of possible worlds it follows that $w_0(p) = 1$ iff $p \in X$ and thus $s_X(p) = 1$. Because the operation $\mathcal{T}$ does not change the value of the background variables, it immediately follows that for background variables $s_X(p) = 1$ iff $s^*_X(p) = 1$. The same way one can show that for background variables $w_0(p) = 0$ iff $s^*_X(p) = 0$.
The induction step. Assume that $\forall q \in \mathcal{P}$ $[H(q) \leq n \Rightarrow (w_0(p) = s^*_X(p))]$. Assume furthermore that we have an arbitrary $p \in \mathcal{P}$ with

$H(p) = n + 1$. Then it follows $p \in \mathcal{P} - B$ with $Z_p = \langle p_1, ..., p_l \rangle$ and for all $1 \leq i \leq l$ it holds $H(p_i) \leq n$. From this we can conclude that for all $1 \leq i \leq l$ it holds $s_X^*(p_i) = w_0(p_i)$ and the value of $f_p(w_0(p_1), ..., w_0(p_l))$ is defined.

case 1 $f_p(w_0(p_1), ..., w_0(p_l)) \neq w_0(p)$. Assume $w_0(p) = 1$. For $p \in \mathcal{P} - B$ we know that in this case $w_0(p) = 1$ iff $p \in X$ iff $s_X(p) = 1$. With definition **??** case (ii.b) it follows $s_X^*(p) = 1$. Thus $w_0(p) = s_X^*(p)$. The same way one shows that the equality holds in case $w_0(p) = 0$.

case 2 $f_p(w_0(p_1), ..., w_0(p_l)) = w_0(p)$. Again, we assume $w_0(p) = 1$. With the assumption of the induction step we can conclude that for all $1 \leq i \leq l$ it holds $w_0(p_i) = s_X^*(p_i)$. Thus, we can conclude that $f_p(s_X^*(p_1), ..., s_X^*(p_l))$ is defined and is evaluated as 1. The case $s_X^*(p) = u$ can be excluded, because otherwise s_X^* would not be a fixed point. We assume that $s_X^*(p) = 0$ and show that this leads to a contradiction. Because $f_p(w_0(p_1), ..., w_0(p_l)) = w_0(p)$ we know that neither p nor $\neg p$ is an element of X. This means $s_X(p) = u$. Because of the monotonicity of $\mathcal{T}$ there is some $1 \leq k$ such that $\mathcal{T}_D^k(s_X)(p) = u$, and $\mathcal{T}^{k+1} + D(s_X)(p) = 0$. Because of definition **??** this means that $f_p(\mathcal{T}_D^k(s_X)(p_1), ..., \mathcal{T}_D^k(s_X)(p_l))$ is defined and evaluated as 0. Thus $f_p(\mathcal{T}_D^k(s_X)(p_1), ..., \mathcal{T}_D^k(s_X)(p_l)) \neq f_p(s_X^*(p_1), ..., s_X^*(p_l))$. This means there is some $1 \leq i \leq l$ with $\mathcal{T}_D^k(s_X)(p_i)$ is defined as either 0 or 1, $s_X^*(p_i)$ is defined as either 0 or 1, but $\mathcal{T}_D^k(s_x)(p_i) \neq s_X^*(p_i)$. This contradicts the monotonicity of $\mathcal{T}$, i.e. fact 2. Hence, we conclude that also in case 2, if $w_0(p) = 1$ then $w_0(p) = s_X^*(p)$. The same way one proves that if $w_0(p) = 0$, then $w_0(p) = s_X^*(p)$.

Add i). The next thing to show is that X is a minimal set with the property $s_X^* = w_0$. Assume $X' \subset X$, i.e. there is some $p \in \mathcal{P}$ with $p \in X$ but $p \notin X'$. By the construction of X we can conclude that either $p \in B$ or $p \in \mathcal{P} - B$ with $Z_p = \langle p_1, ..., p_l \rangle$ and $f_p(w_0(p_1), ..., w_0(p_l)) \neq w_0(p)$. The case $p \in B$ can be excluded immediately, because if $p \notin X'$ then $s_{X'}(p) = u$. Because $\mathcal{T}$ does not change the value of background variables, we conclude that also $s_{X'}^*(p) = u$. But then clearly $s_{X'}^*(p) \neq w_0(p)$. The same reasoning in principle applies also in the second case. Again, we know that $s_{X'}(p) = u$, but, by assumption, $s_{X'}^*(p) = w_0(p) \neq u$. Then from fact 2 we can conclude that there is some $1 \leq k$ with $\mathcal{T}_D^k(s_{X'})(p) = u$ and $\mathcal{T}_D^{k+1}(s_{X'})(p) \neq u$. From definition 3 it follows that $f_p(\mathcal{T}_D^k(s_{X'}(p_1), ..., \mathcal{T}_D^k(s_{X'})(p_l))$ is defined and $f_p(\mathcal{T}_D^k(s_{X'}(p_1), ..., \mathcal{T}_D^k(s_{X'})(p_l)) = \mathcal{T}_D^{k+1}(s_{X'})(p)$. But we also know that $f_p(s^*_{X'}(p_1), ..., s^*_{X'}(p_l)) \neq s^*_{X'}(p)$. That means that either $\mathcal{T}_D^{k+1}(s_{X'})(p) \neq s^*_{X'}(p)$, while both values are either 1 or 0, or there is some $1 \leq i \leq l$ with $s^*_{X'}(p_i) \neq \mathcal{T}_D^k(s_{X'}(p_i)$ while both values are either 0 or 1. Because of the monotonicity of $\mathcal{T}$, i.e. fact 2, both cases are excluded. Hence, we conclude that $s^*_{X'} \neq w_0$. This finishes the proof that X is a basis of w_0.

Part B We have seen that the basis of a possible world always exists. We still have to prove that the basis is uniquely defined. Assume, thus, that there is some set of literals Y that also has the basis property. Following the same strategy we used to prove that X is minimal one can shows that $X \subseteq Y$. Using the fact that Y is minimal it follows $Y = X$. Thus, the basis is uniquely defined.